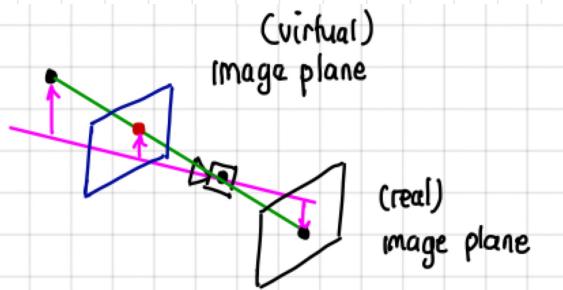
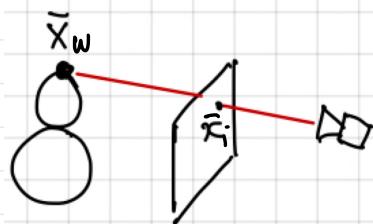


# LECTURE 5: CAMERA GEOMETRY

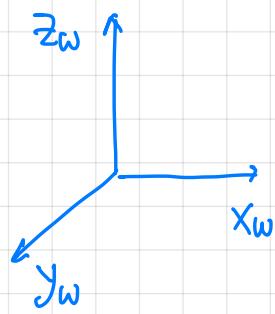
- Geometric Image Formation:

Q: What is the correlation between a 3D point in the world,  $\bar{x}_w = \begin{bmatrix} x_w \\ y_w \\ z_w \end{bmatrix}$  and its corresponding 2D point on an image,  $\bar{x}_i = \begin{bmatrix} x_i \\ y_i \end{bmatrix}$ ?

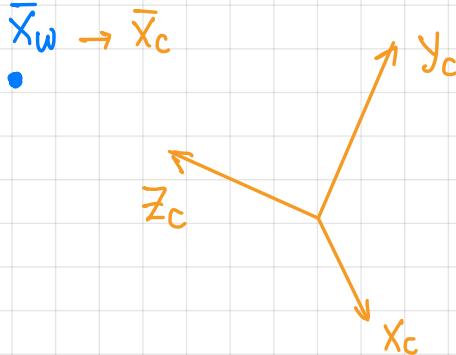


Coordinate Systems

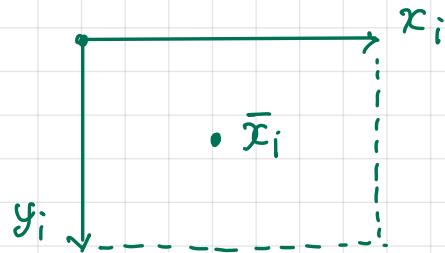
① World



② Camera



③ Image



$$\bar{x}_i = P \bar{x}_w$$

$3 \times 1$     $3 \times 4$     $4 \times 1$

$\Rightarrow$  What is  $P$ ?

$$\begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix} = P \begin{bmatrix} x_w \\ y_w \\ z_w \end{bmatrix}$$

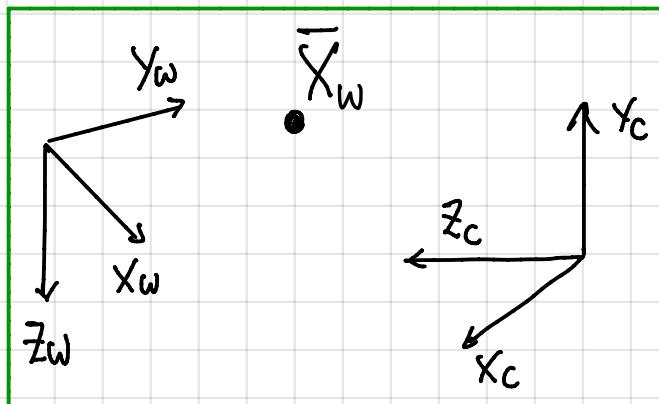
# [●] Camera Matrix ( $\mathbb{P}$ )

$$\bar{\mathbf{x}}_i = \mathbb{P} \bar{\mathbf{x}}_w$$

$3 \times 1 \quad 3 \times 4 \quad 4 \times 1$

What matrix is  $\mathbb{P}$ ?

(1) From the world to the camera:



$$\bar{\mathbf{x}}_c = \mathbb{T} \mathbb{R} \bar{\mathbf{x}}_w$$

$\bar{\mathbf{x}}_w$  = the coordinates of  $\bar{\mathbf{x}}$  in the world coordinates.

$\bar{\mathbf{x}}_c$  = the coordinates of  $\bar{\mathbf{x}}$  in the camera coordinates.

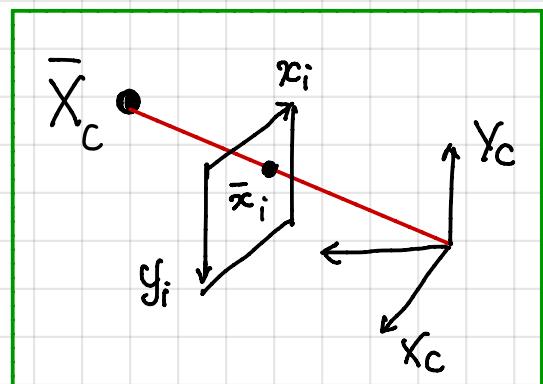
$\mathbb{R}$  = a rotation matrix that rotates the world coordinates of  $\bar{\mathbf{x}}$  to the camera coordinates.

$\mathbb{T}$  = a translation matrix that translates the world coordinates of  $\bar{\mathbf{x}}_w$  to the camera coordinates



What matrices are  $\mathbb{R}$  &  $\mathbb{T}$  exactly?

(2) From the camera to the image:



$$\bar{\mathbf{x}}_i = \mathbb{K} \bar{\mathbf{x}}_c$$

$\mathbb{K}$ : ① to scale down  
② to translate from the center of the image to the top left of the image

$$\mathbb{K} = \begin{bmatrix} 1 & 0 & x_0 \\ 0 & 1 & y_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} f_x & 0 & 0 \\ 0 & f_y & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

to translate

to scale down

$$\mathbb{K} = \begin{bmatrix} f_x & 0 & x_0 \\ 0 & f_y & y_0 \\ 0 & 0 & 1 \end{bmatrix}$$

the center of the image

$f_x \approx f_y$

$$\bar{X}_c = R \bar{T} \bar{X}_w$$

where :

$$\bar{T} = \begin{bmatrix} 1 & 0 & 0 & T_x \\ 0 & 1 & 0 & T_y \\ 0 & 0 & 1 & T_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad ; \quad R = \begin{bmatrix} R_{11} & R_{12} & R_{13} & 0 \\ R_{21} & R_{22} & R_{23} & 0 \\ R_{31} & R_{32} & R_{33} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

We can write :  $\bar{T} R = \underbrace{\begin{bmatrix} R_{11} & R_{12} & R_{13} & T_x \\ R_{21} & R_{22} & R_{23} & T_y \\ R_{31} & R_{32} & R_{33} & T_z \\ 0 & 0 & 0 & 1 \end{bmatrix}}_{4 \times 4} = \begin{bmatrix} R & | & \bar{T} \\ 0 & | & 1 \end{bmatrix}$

$$\bar{X}_c = \underbrace{\bar{T}}_{4 \times 1} \underbrace{R}_{4 \times 4} \underbrace{\bar{X}_w}_{4 \times 1} = \begin{bmatrix} R & | & \bar{T} \\ 0 & | & 1 \end{bmatrix} \bar{X}_w$$

$$\bar{X}_c = \underbrace{\begin{bmatrix} R & | & \bar{T} \end{bmatrix}}_{3 \times 4} \underbrace{\bar{X}_w}_{4 \times 1}$$

$\bar{X}_c$  is in the inhomogeneous coordinates  
 $\bar{X}_w$  is in the homogeneous coordinates



Hence :

$$\begin{aligned} \bar{X}_i &= \underbrace{P}_{3 \times 1} \bar{X}_w = \underbrace{K}_{3 \times 3} \underbrace{\bar{X}_c}_{4 \times 1} \\ &= \underbrace{K}_{3 \times 3} \underbrace{\begin{bmatrix} R & | & \bar{T} \end{bmatrix}}_{3 \times 4} \underbrace{\bar{X}_w}_{4 \times 1} \end{aligned}$$

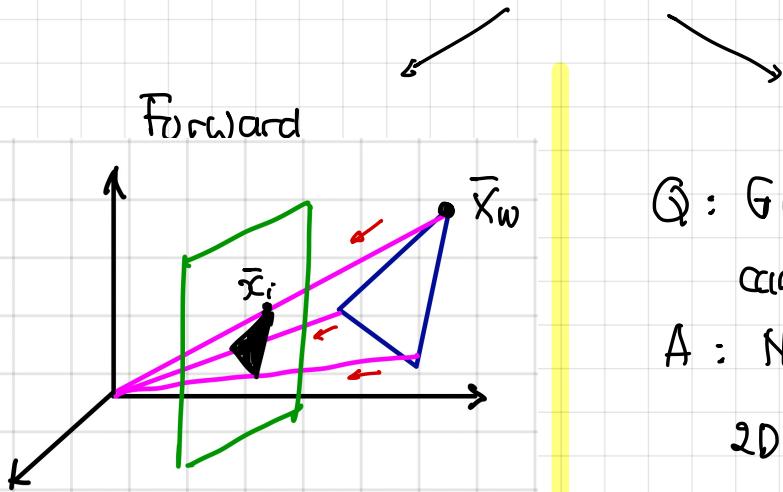
Intrinsic parameters      Extrinsic parameters  
 (Internal camera parameters)      (External camera parameters)



Camera geometric calibration / camera resectioning :

To estimate the values of  $K, R, \bar{T}$ .

## [•] Forward & Backward Projection



$$\bar{x}_i = \bar{P} \bar{X}_w$$

Note:  $\bar{x} = \begin{bmatrix} x_i \\ y_i \\ w \end{bmatrix}$

Hence, don't forget to normalize:  $\bar{x} = \begin{bmatrix} x_i/w \\ y_i/w \\ 1 \end{bmatrix}$

in World coordinates:

camera coordinates:

$$\bar{C}_c = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

$$[\bar{R} \mid \bar{t}] \bar{C}_w = \bar{C}_c = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

$$\bar{I} \bar{R} \bar{C}_w = \bar{0}$$

$$\bar{R} \bar{C}_w + \bar{t} = \bar{0}$$

$$\bar{R} \bar{C}_w = -\bar{t}$$

$$\bar{C}_w = -\bar{R}^{-1} \bar{t} = -\bar{R}^T \bar{t}$$

Q: Given a point  $\bar{x}_i$  on an image &  $\bar{P}$ , can we recover  $\bar{X}_w$ ?

A: No. Backward projection can only do:  
2D image      3D world

point → line

line → plane

plane → cone / volume

Backward projection: To find a set of points forming a ray passing through the camera center,  $\bar{C}_w$ , &  $\bar{x}_i$  in the world coordinates

$\bar{x}_i$  in World coordinates:

$$\bar{x}_i = \bar{P} \bar{X}_w$$

$3 \times 1$        $3 \times 4$        $4 \times 1$

$$\hat{\bar{X}}_w = \bar{P}^+ \bar{x}_i$$

$4 \times 1$        $4 \times 3$        $3 \times 1$

$$\bar{P}^+ = \frac{(\bar{P}^T \bar{P})^{-1}}{4 \times 4} \bar{P}^T$$

note:  $\hat{\bar{X}}_w \neq \bar{X}_w$

but both are on the same line

The line can be drawn from

$\bar{C}_w$  &  $\hat{\bar{X}}_w$ .

under determined system  
many possible solutions

## [•] Camera Calibration

Problem statement : Given  $N$  corresponding points  $\bar{x}_n \leftrightarrow \bar{X}_n$   
 estimate  $K, R, t$  such that  $\bar{x}_n = P \bar{X}_n$ .

Solution :

① Compute  $P$

② Decompose  $P$  into  $K, R$ , and  $t$



[6] Step 1 : Compute  $P$

$$\bar{x}_i = \begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix} = \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \end{bmatrix} \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix}$$

$$x_i = \frac{p_{11} x_w + p_{12} y_w + p_{13} z_w + p_{14}}{p_{31} x_w + p_{32} y_w + p_{33} z_w + p_{34}}, \quad y_i = \frac{p_{21} x_w + p_{22} y_w + p_{23} z_w + p_{24}}{p_{31} x_w + p_{32} y_w + p_{33} z_w + p_{34}}$$

$$\Leftrightarrow x_i(p_{31} x_w + p_{32} y_w + p_{33} z_w + p_{34}) = p_{11} x_w + p_{12} y_w + p_{13} z_w + p_{14}$$

$$y_i(p_{31} x_w + p_{32} y_w + p_{33} z_w + p_{34}) = p_{21} x_w + p_{22} y_w + p_{23} z_w + p_{24}$$

$$p_{11} x_w + p_{12} y_w + p_{13} z_w + p_{14} + \emptyset p_{21} + \emptyset p_{22} + \emptyset p_{23} + \emptyset p_{24} - x_i x_w p_{31} - x_i y_w p_{32} - x_i z_w p_{33} - x_i p_{34} = \emptyset$$

$$p_{11} \emptyset + p_{12} \emptyset + p_{13} \emptyset + p_{14} \emptyset + x_w p_{21} + y_w p_{22} + z_w p_{24} - y_i x_w p_{31} - y_i y_w p_{32} - y_i z_w p_{33} - y_i p_{34} = \emptyset$$

Vectorization :  $\begin{bmatrix} x_w & y_w & z_w & | & \emptyset & \emptyset & \emptyset & \emptyset & -x_i x_w & -x_i y_w & -x_i z_w & -x_i \\ \emptyset & \emptyset & \emptyset & | & x_w & y_w & z_w & | & -y_i x_w & -y_i y_w & -y_i z_w & -y_i \end{bmatrix} \bar{P} = 0$

where :  $\bar{P} = [p_{11} \ p_{12} \ p_{13} \ p_{14} \ p_{21} \ p_{22} \ p_{23} \ p_{24} \ p_{31} \ p_{32} \ p_{33} \ p_{34}]^T$

Thus :  $A \bar{P} = \bar{0}$  ;  $\bar{P} \neq 0$  ; otherwise we end up with  
 $2 \times 2 \quad 2 \times 1 \quad 2 \times 1$  a trivial solution.

## Step 1 (Compute $\bar{P}$ ): Homogeneous Linear System

$A\bar{P} = 0 \rightarrow$  Since 1 pair of  $(\bar{x}_n, \hat{x}_n)$  provides 2 equations,  
 $2 \times 12 \quad 12 \times 1$   
to estimate  $\bar{P}$ , we need at least 6 pairs ( $N \geq 6$ ):

$$A\bar{P} = 0 \Rightarrow \text{Homogeneous Linear System}$$

$(2N \times 12) \quad (12 \times 1)$



How to solve this?

(A) Estimate the initial value of  $\bar{p}$  using SVD

(B) Refine  $\bar{p}$  using least-squares

$$\text{minimize } \|A\bar{p}\|$$

$$\text{subject to } \|\bar{p}\| = 1$$

$\|A\bar{p}\|$  means

the magnitude of  $A\bar{p}$ .

This is a constraint to avoid the trivial solution of  $\bar{p} = 0$ .

$$A = UDV^T$$

$\bar{p}$  is the last row of  $V^T$ .

# Step 1.B : Refine $\bar{P}$ using Least Squares

$$\bar{P}^* = \underset{\{\bar{P}\}}{\operatorname{argmin}} E(\bar{P}) = \underset{\{\bar{P}\}}{\operatorname{argmin}} \frac{1}{2} \sum_n^N \left[ \bar{x}_n - \bar{P} \bar{X}_n \right]^2$$

Note:  $\bar{a}^2$  doesn't exist, squaring a vector:  $\bar{a}^T \bar{a}$

The expression of  $\bar{a}^2$  is only for writing convenience.

Hence:

$$E(P) = E(\bar{P}) = \frac{1}{2} \sum_n^N$$

$$\begin{cases} x_i - \frac{p_{11}x_w - p_{12}y_w - p_{13}z_w - p_{14}}{p_{21}x_w - p_{22}y_w - p_{23}z_w - p_{24}} \\ y_i - \frac{p_{21}x_w - p_{22}y_w - p_{23}z_w - p_{24}}{p_{31}x_w - p_{32}y_w - p_{33}z_w - p_{34}} \end{cases}^2$$

$$E(\bar{P}) = \frac{1}{2} \sum_n^N \left[ A_n \bar{P} \right]^2$$

Minimization:

$$\frac{d E(\bar{P})}{d \bar{P}} = 0 \quad ; \quad \begin{matrix} 12 \times 1 \\ 12 \times 1 \end{matrix}$$

$$\frac{d E}{d \bar{P}} = \sum_n^N \begin{matrix} 12 \times 2 \\ 12 \times 1 \end{matrix} [A_n^T \quad A_n] \begin{matrix} 12 \times 2 \\ 12 \times 1 \end{matrix} \bar{P}$$

$$= \sum_n^N (A_n^T A_n) \bar{P} \quad ; \quad \bar{P} \neq 0$$

Since we can't have a closed-form solution for  $\bar{P}$ , we use Gradient Descent or Newton's methods:

$$\text{Gradient Descent: } \bar{P}^{\text{new}} = \bar{P}^{\text{old}} + \alpha \frac{d E}{d \bar{P}} \Big|_{\bar{P} = \bar{P}^{\text{old}}} \quad \begin{matrix} 12 \times 1 \\ 12 \times 1 \end{matrix}$$

$$\text{In the beginning: } \bar{P}^{\text{init}} = \bar{P}^{\text{SVD}}$$

$$\text{Newton's method: } \alpha = \left( \frac{d^2 E}{d \bar{P}^2} \Big|_{\bar{P} = \bar{P}^{\text{old}}} \right)^{-1} \quad ; \quad \begin{matrix} \text{Inverse of} \\ \text{second derivative} \end{matrix}$$

## Step 2 : Decompose $P$ to $K, R$ and $\bar{t}$

Steps:

1. Extract  $M$  (a  $3 \times 3$  matrix) from the first  $3 \times 3$  submatrix of  $P$ .
2. Factor  $M$  into  $K R$  using  $RQ$  decomposition

$$M = KR \quad ; \text{ recall } K \text{ is a triangular matrix}$$



Example :

$$A = QR \rightarrow \begin{bmatrix} 2 & 2 \\ 1 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} -0.89 & -0.45 \\ -0.45 & 0.89 \end{bmatrix} \begin{bmatrix} -2.24 & -2.24 \\ 0 & 0 \end{bmatrix}$$

$$A = RQ \rightarrow \begin{bmatrix} 2 & 2 \\ 1 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 0 & -2.83 \\ 0 & -1.41 \end{bmatrix} \begin{bmatrix} -0.71 & 0.71 \\ -0.71 & -0.71 \end{bmatrix}$$

$$3. \bar{t} = K^{-1} \begin{pmatrix} p_{14} \\ p_{24} \\ p_{34} \end{pmatrix} \rightarrow \text{Why?}$$



Because:  $K^{-1} P = [R | \bar{t}]$

$$K^{-1} \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \end{bmatrix} = \begin{bmatrix} R_{11} & R_{12} & R_{13} & \bar{t}_x \\ R_{21} & R_{22} & R_{23} & \bar{t}_y \\ R_{31} & R_{32} & R_{33} & \bar{t}_z \end{bmatrix}$$

## [•] Intrinsic Camera Properties, $\mathbb{K}$ :

### Skew Parameters

Previously we define:  $\mathbb{K} = \begin{bmatrix} f_x & 0 & x_0 \\ 0 & f_y & y_0 \\ 0 & 0 & 1 \end{bmatrix}$

A more common definition:

$$\mathbb{K} = \begin{bmatrix} f_x & s & x_0 \\ 0 & f_y & y_0 \\ 0 & 0 & 1 \end{bmatrix}$$

where :

$s$  = the skew parameter

Three operations in  $\mathbb{K}$ :

$$\mathbb{K} = \begin{bmatrix} 1 & 0 & x_0 \\ 0 & 1 & y_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} f_x & 0 & 0 \\ 0 & f_y & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & s & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

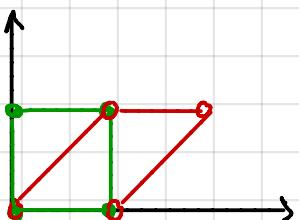
translation

scaling

shearing

Shearing example:

$$\begin{bmatrix} 1 & \textcolor{green}{s} & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 1 & 2 \\ 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$



- Notes:
1. In most cameras,  $s=0$
  2. Why does the skew happen horizontally? It's rare that images are skewed vertically.

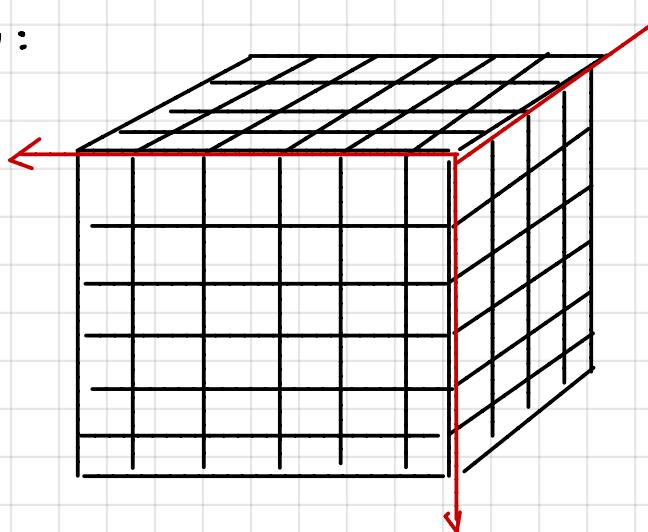
## [0] Practical Notes on Calibration

1. Q: How to obtain the corresponding points:  $\bar{x}_n$  and  $\tilde{x}_n$ ?

A: Use either a 2D planar surface or a 3D box with identifiable patterns (= a checker pattern)



3D calibration box:



2. Q: Why  $\bar{c} = -R^T \bar{t}$ ?

A: We know that:  $[R | t] \bar{c} = 0$ ,

which means when we transform  $\bar{c}$ , the center of the camera in the world coordinate system, onto the camera coordinate system, it should be at the origin,  $\bar{0}$ .

$$[R | t] \bar{c} = 0$$

$4 \times 3 \quad 4 \times 1 \quad 4 \times 1$

$$; \quad R = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \\ 0 & 0 & 0 \end{bmatrix}$$

$$R \bar{c} + \bar{t} = 0$$

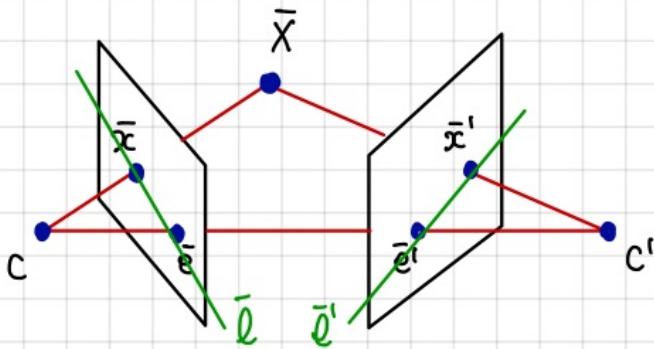
$$; \quad R = \begin{bmatrix} r_{11} & r_{12} & r_{13} & 0 \\ r_{21} & r_{22} & r_{23} & 0 \\ r_{31} & r_{32} & r_{33} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$R \bar{c} = -\bar{t}$$

$$\begin{aligned} \bar{c} &= -R^{-1} \bar{t} \\ &= -R^T \bar{t} \end{aligned}$$

} since  $R$  is an orthonormal matrix.

## [•] Fundamental Matrix



$c$  = the camera center  
 $\bar{e}$  = the epipole  
 $\bar{l}$  = the epipolar line

What are the possible correlations between  $\bar{x}$  &  $\bar{x}'$  above?

Homography

Fundamental Matrix

$$\bar{x}' = H \bar{x}$$

- (1)  $\bar{x}'^T F \bar{x} = 0$
- (2)  $\bar{l}' = F \bar{x}$

What is  $F$ ?

Two points create a line.

$$\bar{l}' = \bar{x}' \times \bar{e}'$$

$$\bar{l}' = [e']_x \bar{x}'$$

Skew-symmetric matrix:

$$[e']_x = \begin{bmatrix} 0 & -e'_z & e'_y \\ e'_z & 0 & -e'_x \\ -e'_y & e'_x & 0 \end{bmatrix}$$

↓

Purpose:

to transform a cross product with a vector to a matrix multiplication

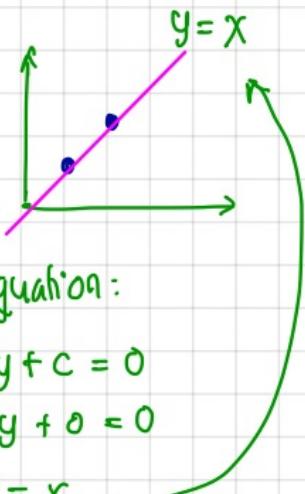
e.g.:  $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \times \begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix}$

$$= \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$$

↓  $\bar{l}$

line equation:  
 $ax + by + c = 0$   
 $-1x + 1y + 0 = 0$

$y = x$



$$\bar{x}' = H \bar{x}$$

$$\bar{l}' = [e']_x H \bar{x}$$

$\bar{l}' = F \bar{x}$ ; where  $F = [e']_x H$

Another look at  $\bar{F}$ :

$$\bar{x} = \bar{P} \bar{X}$$

Using the backward projection:

$$\bar{X} = \bar{C} + \lambda P^+ \bar{x}$$

This is a ray governed by  
two points:  $P^+ \bar{x}$  &  $\bar{C}$

Projecting the two points ( $P^+ \bar{x}$  and  $\bar{C}$ ) onto another image:

point  $P^+ \bar{x} \rightarrow P' P^+ \bar{x}$

point  $\bar{C} \rightarrow P' \bar{C}$

$$\rightarrow P' \bar{C} = \bar{e}'$$

The epipolar line of these two projected points:

$$\begin{aligned}\bar{L}' &= (P' P^+ \bar{x}) \times (P' \bar{C}) = (P' P^+ \bar{x}) \times \bar{e}' \\ &= [\bar{e}']_x P' P^+ \bar{x}\end{aligned}$$

Thus, we can also define:  $\bar{F} = [\bar{e}']_x P' P^+$

$$\boxed{\bar{F} = [P' \bar{C}]_x P' P^+}$$

From the last equation, we can infer that:

1.  $\bar{F}$  is independent from the world structures ( $X$ ), and depends only on the camera properties.
2. An image pair has a single value of  $\bar{F}$ .

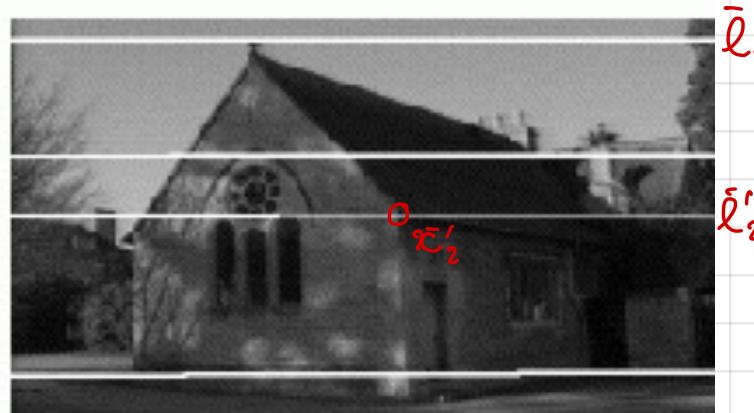
## Properties of F

- $F$  is a rank 2 homogeneous matrix with 7 degrees of freedom.
- **Point correspondence:** If  $x$  and  $x'$  are corresponding image points, then  $x'^\top Fx = 0$ .
- **Epipolar lines:**
  - ◊  $l' = Fx$  is the epipolar line corresponding to  $x$ .
  - ◊  $l = F^\top x'$  is the epipolar line corresponding to  $x'$ .
- **Epipoles:**
  - ◊  $Fe = 0$        $F^\top e' = 0$
- **Computation from camera matrices  $P, P'$ :**
  - ◊  $F = [P'c]_\times P'P^+$ , where  $P^+$  is the pseudo-inverse of  $P$ , and  $c$  is the centre of the first camera. Note,  $e' = P'c$ .
  - ◊ Canonical cameras,  $P = [I \mid o]$ ,  $P' = [M \mid m]$ ,  
 $F = [e']_\times M = M^{-\top}[e]_\times$ , where  $e' = m$  and  $e = M^{-1}m$ .

The lines are the epipolar lines corresponding to the points on the left figure.



left



Right